ZooKeys 319: 363–377 (2013)
doi: 10.3897/zookeys.319.4178
www.zookeys.org

RESEARCH ARTICLE

A peer-reviewed open-access journal
ZooKeys
Launched to accelerate biodiversity research

Secondary structure models of 18S and 28S rRNAs of the true bugs based on complete rDNA sequences of *Eurydema maracandica* Oshanin, 1871 (Heteroptera, Pentatomidae)

Shasha Yu[1], Yanhui Wang[1], Dávid Rédei[1], Qiang Xie[1], Wenjun Bu[1]

1 *Institute of Entomology, College of Life Sciences, Nankai University, 94 Weijin Road, Tianjin 300071, China*

Corresponding author: *Qiang Xie* (qiangxie@nankai.edu.cn); *Wenjun Bu* (wenjunbu@nankai.edu.cn)

Academic editor: *S. Grozeva* | Received 25 November 2012 | Accepted 7 January 2013 | Published 30 July 2013

Citation: Yu S, Wang Y, Rédei D, Xie Q, Bu W (2013) Secondary structure models of 18S and 28S rRNAs of the true bugs based on complete rDNA sequences of *Eurydema maracandica* Oshanin, 1871 (Heteroptera, Pentatomidae). In: Popov A, Grozeva S, Simov N, Tasheva E (Eds) Advances in Hemipterology. ZooKeys 319: 363–377. doi: 10.3897/zookeys.319.4178

Abstract

The sequences of 18S and 28S rDNAs have been used as molecular markers to resolve phylogenetic relationships of Heteroptera for two decades. The complete sequences of 18S rDNAs have been used in many studies, while in most studies only partial sequences of 28S rDNAs have been used due to technical difficulties of amplifying the complete lengths. In this study, we amplified the complete 18S and 28S rDNA sequences of *Eurydema maracandica* Oshanin, 1871, and reconstructed the secondary structure models of the corresponding rRNAs. In addition, and more importantly, all of the length variable regions of 18S rRNA were compared among 37 families of Heteroptera based on 140 sequences, and the D3 region of 28S rRNA was compared among 51 families based on 84 sequences. It was found that 8 length variable regions could potentially serve as molecular synapomorphies for some monophyletic groups. Therefore discoveries of more molecular synapomorphies for specific clades can be anticipated from amplification of complete 18S and 28S rDNAs of more representatives of Heteroptera.

Keywords

rRNA, secondary structure, Heteroptera, molecular synapomorphy, *Eurydema maracandica*

Copyright *Shasha Yu et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC-BY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Introduction

Each cluster of rDNA in turn contains external transcribed spacer (ETS), 18S rDNA, internal transcribed spacer 1 (ITS1), 5.8S rDNA, ITS2 and 28S rDNA (Hillis, Dixon 1991). These genes and spacer regions, separately or in combination, have been used as molecular markers for resolving phylogenetic relationships of Heteroptera for two decades from Wheeler et al. (1993). Among them, the length-conservative regions of 18S rDNA and 28S rDNA have been mostly used for resolving phylogenetic relationships between higher level groups (Wheeler et al. 1993, Xie et al. 2005, Schuh et al. 2009), while the length-variable regions (LVRs), ITS1 and ITS2 have always been used at species level (Marcilla et al. 2001, Hebsgaard et al. 2004, Damgaard et al. 2005, Calleros et al. 2010).

As far as the rDNAs are concerned, complete sequences of 18S rDNA have been used in various phylogenetic studies on Heteroptera (Tian et al. 2008, Jung, Lee 2011). Comparatively, so far only partial sequences of 28S rDNAs have been involved in phylogenetic studies on Heteroptera due to technical difficulties of amplifying the complete lengths. These difficulties are twofold. First, the amplification of 28S rDNA is more likely interfered by the hairpin structures, or tandem replicates of single nucleotides or oligonucleotides. Second, universal primers are unavailable for amplifying the 28S rDNA of Heteroptera.

Reconstruction of the complete secondary structure model of 28S rRNA can only be achieved by full sequencing of 28S rDNA. According to the results of comparative studies of 18S rRNAs of Insecta (Xie et al. 2009) and Hemiptera (Xie et al. 2008), the secondary structure model can have significance in two aspects. First, the alignment results of rDNAs can be corrected based on the secondary structure models of the corresponding rRNAs (Kjer et al. 2006). Further, it is possible to evaluate the potential influences to the results of phylogenetic reconstruction. Second, the secondary structure model provides a novel view and datasets to discover synapomorphies, such as group-specific insertions or deletions (indels), exactly the same length of some LVRs, and length expansions. In this study, the secondary structure models of 18S and 28S rRNAs of Heteroptera were reconstructed based on corresponding rDNA sequences of the pentatomid species *Eurydema maracandica* Oshanin, 1871. Additionally, all LVRs of complete 18S rRNA and the Divergent region 3 (D3) of 28S rRNA were compared based on data of nearly all available families of Heteroptera from GenBank (www.ncbi.nlm.nih.gov/Genbank), with exception of a few problematic sequences.

Methods

Taxon sampling

According to the currently accepted classification of Heteroptera (Schuh and Slater 1995, Schuh et al. 2008), heteropterans are divided into 7 infraorders, 23 superfami-

lies and 76 families. In this study, the 18S rDNA dataset comprises 140 sequences representing 7 infraorders, 16 superfamilies and 37 families of Heteroptera. The 28S rDNA dataset, which only included the D3 region, comprises 84 sequences of 7 infraorders, 21 superfamilies and 51 families.

Molecular experiments

The specimen was collected from Yining, Xinjiang Uyghur Autonomous Region, China (43°56'N, 81°19'E, 570m) on July 26, 2011 by Qiang Xie. In the alcohol-kept insects' collection of Nankai University, the ID number for the voucher specimen is NKU0150177. The species was identified based on the available literature and comparative material, its identity was confirmed by P. Kment. The specimen perfectly fits the original description (Oshanin 1871) and the available detailed redescriptions and illustrations (Stichel 1961, Putshkov 1965, Zheng 1983, Nonnaizab 1988) of *E. maracandica*.

Genomic DNA was extracted from thoracic tissue of ethanol-preserved specimen. Total genomic DNA was isolated using the CTAB-based method (Reineke et al. 2002). The primers Ns1 and Ns8 used for amplifying the 18S rDNA were those used by Barker et al. (2003). The others were designed by the software Primer Premier 5.00 (Lalitha 2000). These primer sets were used to amplify six overlapping fragments of 28S rDNA and two overlapping fragments of 18S rDNA. All of the sequences were sequenced in both directions. In addition, 18S rDNA was cloned, which used pEASY-T3 (TransGen, Beijing, China) as vector, following the manufacturer's instructions. Amplification was carried out in a 50 μL volume reaction, with 1.5 units of *LA Taq* DNA Polymerase, 2.5 mM of dNTP and 10 μM of each primer. The thermal cycling program of PCR consisted of an initial denaturation at 94°C for 1 min, followed by 35 cycles (94°C for 30 s, 49–57°C for 30 s, 72°C for 1 min), and ending with a final extension at 72°C for 8 min. Colony PCR was carried out in a 25 μL volume reaction, and consisted of 30 amplification cycles. All of the primer pairs and annealing temperatures are listed in Supplementary File 1. The GenBank accession numbers of 18S and 28S rDNAs of *Eurydema maracandica* are JX997807 and JX997806 respectively.

Reconstruction of secondary structure and Sequence alignment

The secondary structure of 18S rRNA was slightly revised from the universal model of insects (Xie et al. 2009) based on the sequence of *Drosophila melanogaster* (GenBank accession number M21017). In the group specific optimization of the universal model according to the sequences of Heteroptera, secondary structures of LVRs were reconstructed using the thermodynamics folding method of the software RNAstructure 5.4 (Reuter and Mathews 2010). The secondary structure of 28S rRNA was revised from the model of insects (Wang et al. 2013) based on the sequence of *Drosophila mela-*

nogaster (GenBank accession number M21017). The numbering system for LVRs of 28S rRNA followed the D system (Hassouna et al. 1984, Ellis et al. 1986). In the optimization of the previous model of insects, secondary structures of LVRs were reconstructed using RNAstructure 5.4 as well. The re-calculated results in this study were selected under the principle of co-variation: the fewer the secondary structural elements, especially the paired regions, are destroyed by each sequence, the better the model is (Gutell et al. 2002). We simplified the principle in this study as: the longer the stems are kept by each sequence, the better the model is. Sequence alignment was performed by ClustalW, which is imbedded within BioEdit 7.1.3 (Hall 2011). Manual alignments were based on the rRNA secondary structure models reconstructed in this study.

Data resources

The data underpinning the analyses reported in this paper are deposited in the Dryad Data Repository at doi: 10.5061/dryad.j8kp5

Results and discussion

General description of secondary structures

The complete sequence of 18S rDNA of *Eurydema maracandica* is 1,894 bp. In the secondary structure of the corresponding rRNA (Fig. 1) the LVRs L and X regions were specifically optimized. The 18S rDNA sequence of *Eurydema maracandica* was aligned with other 140 orthologous sequences of 37 families in Heteroptera. The alignment result suggested that the local length variations of 18S rDNA are localized in 13 independent LVRs except for various indels. Most LVRs were restrained in three domains, which were previously named as V2, V4 and V7 (Neefs et al. 1993). Among all of the 13 LVRs, LVR L is the most variable one in length.

The complete sequence of 28S rDNA of *Eurydema maracandica* is 4,030 bp. In the secondary structure of corresponding rRNA (Figs 2, 3), the LVRs D2, D3, D7, D8, D10, and D11 were specifically optimized. Compared with the secondary structure of the universal model of insect 28S rRNA, LVRs of heteropteran 28S rRNAs are distributed in 10 regions (D2–D11). Only D3 region of most superfamilies has corresponding data in GenBank, therefore only this region was comparatively analyzed and the results may serve as an inspiring case for the other LVRs of 28S rRNA. Alignment results suggested that in the D3 region, LVRs can be further divided into three sections (D3-1, D3-2 and D3-3), which are interspaced by several short length-conservative regions. Among these three separate LVRs, D3-2 is the most variable one in length.

According to the above results, the local length variations of heteropteran rRNAs are moderate in comparison with the length variations of eukaryotic rRNAs (Xie et

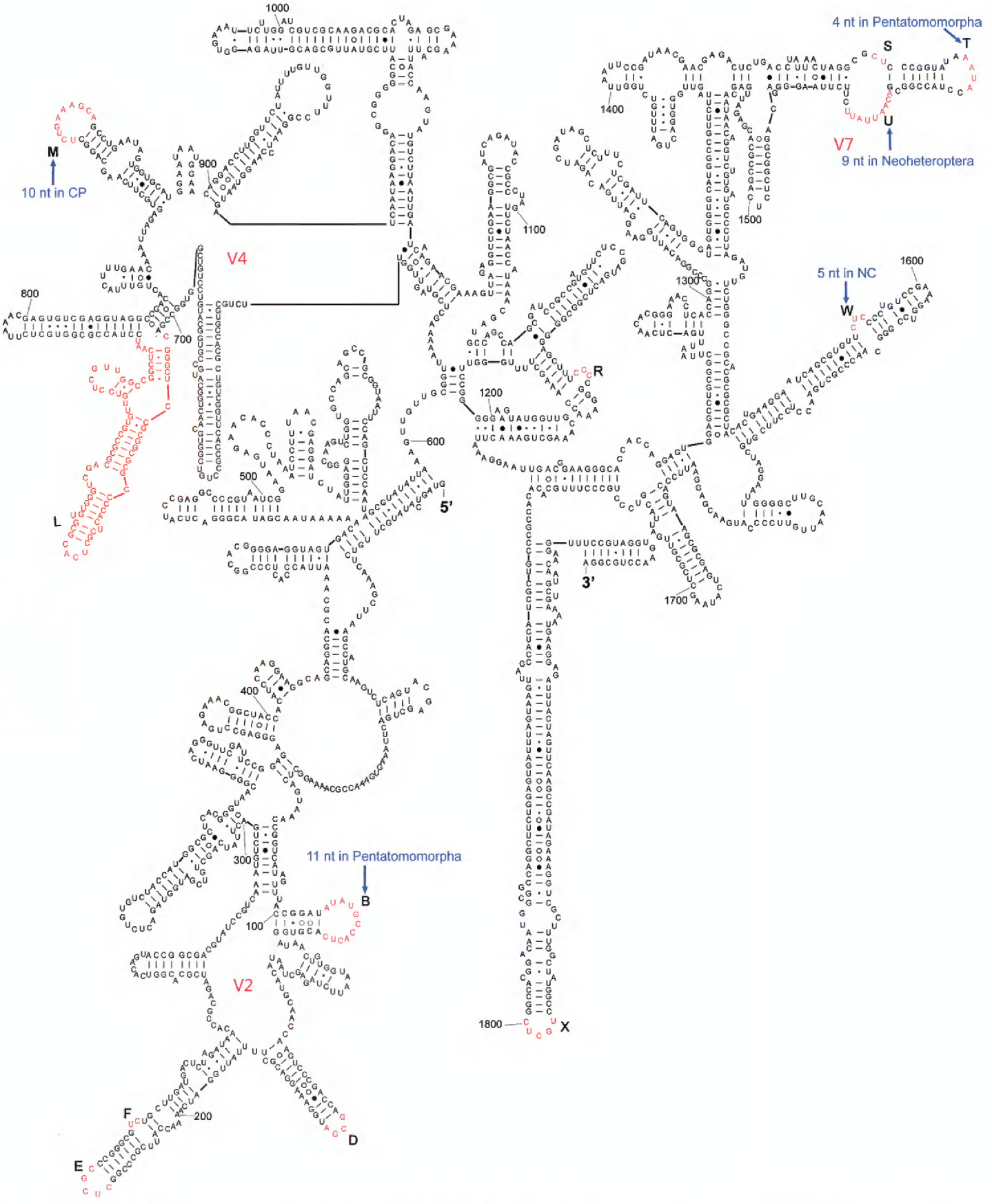

Figure 1. Secondary structure model of 18S rRNA of *Eurydema maracandica*. The bases marked in black represent length-conservative regions, and the bases labeled as capital letters **B** to **W** in red represent 13 LVRs. CP and NC represent monophyletic groups Cimicomorpha+Pentatomomorpha and Naboidea+Cimicoidea, respectively. Base pairing is indicated as follows: standard canonical pairs by lines (G–C, A–U), wobble G:U pairs by dots (G·U), A:G or A:C pairs by open circles (A○G, A○C), and other non-canonical pairs by filled circles (e.g., A●A).

al. 2011). The local length variations of heteropteran rRNAs did not cause overlap between adjacent regions (See Supplementary Files 2, 3) like it was documented in the suborder Sternorrhyncha of Hemiptera (Xie et al. 2008) and some other insect orders

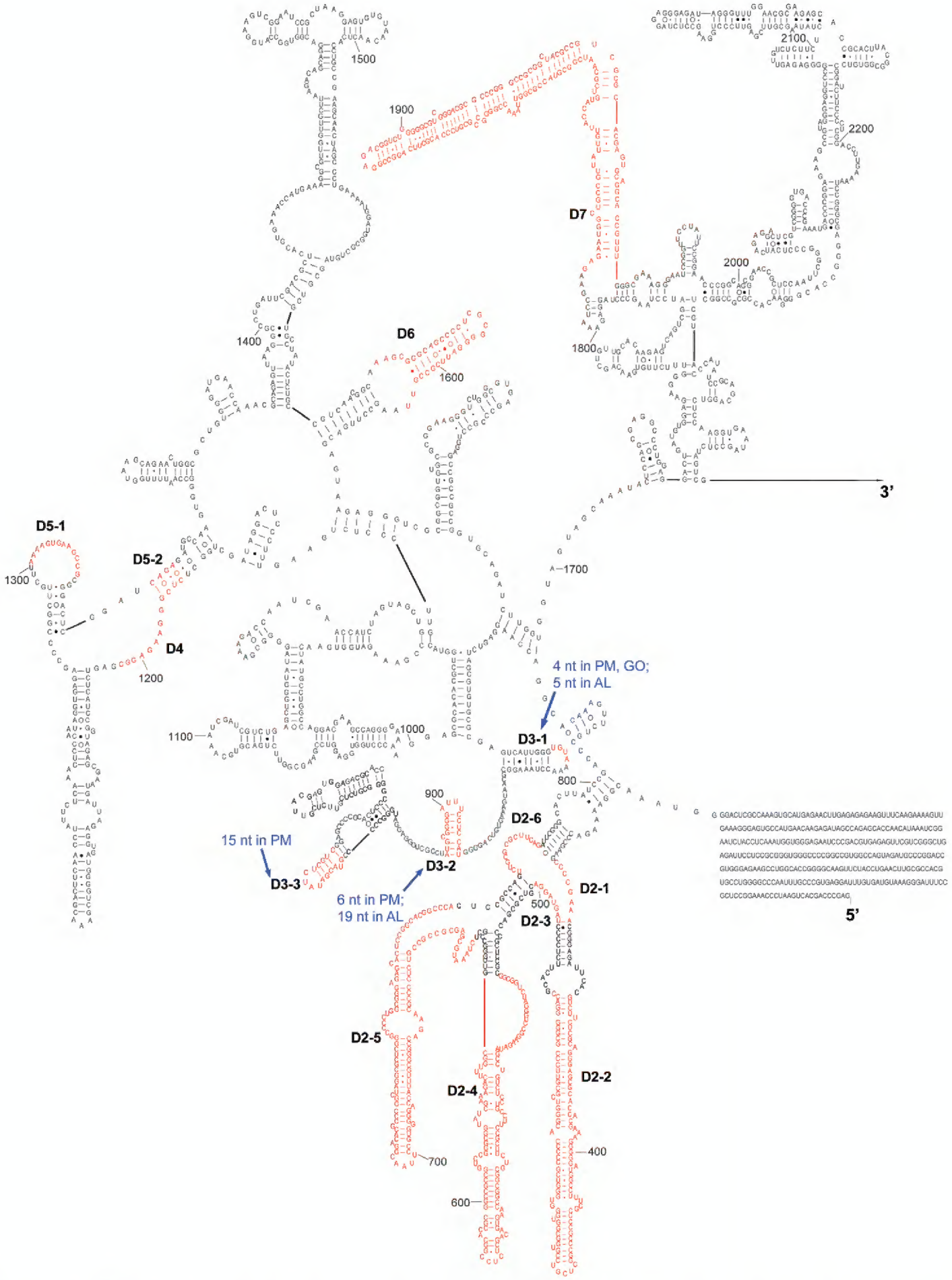

Figure 2. The 5'-half part of secondary structure model of 28S rRNA of *Eurydema maracandica*. The numbers D2 to D7 represent six LVRs. PM, GO and AL represent monophyletic groups Paraphrynovelii dae+Macroveliidae, Gelastocoridae+Ochteridae and Acanthosomatidae+Lestoniidae, respectively.

(Xie et al. 2009). However, the ambiguity in alignment caused by length variations might have impact on the results of phylogenetic reconstruction of Heteroptera, and it deserves further evaluation.

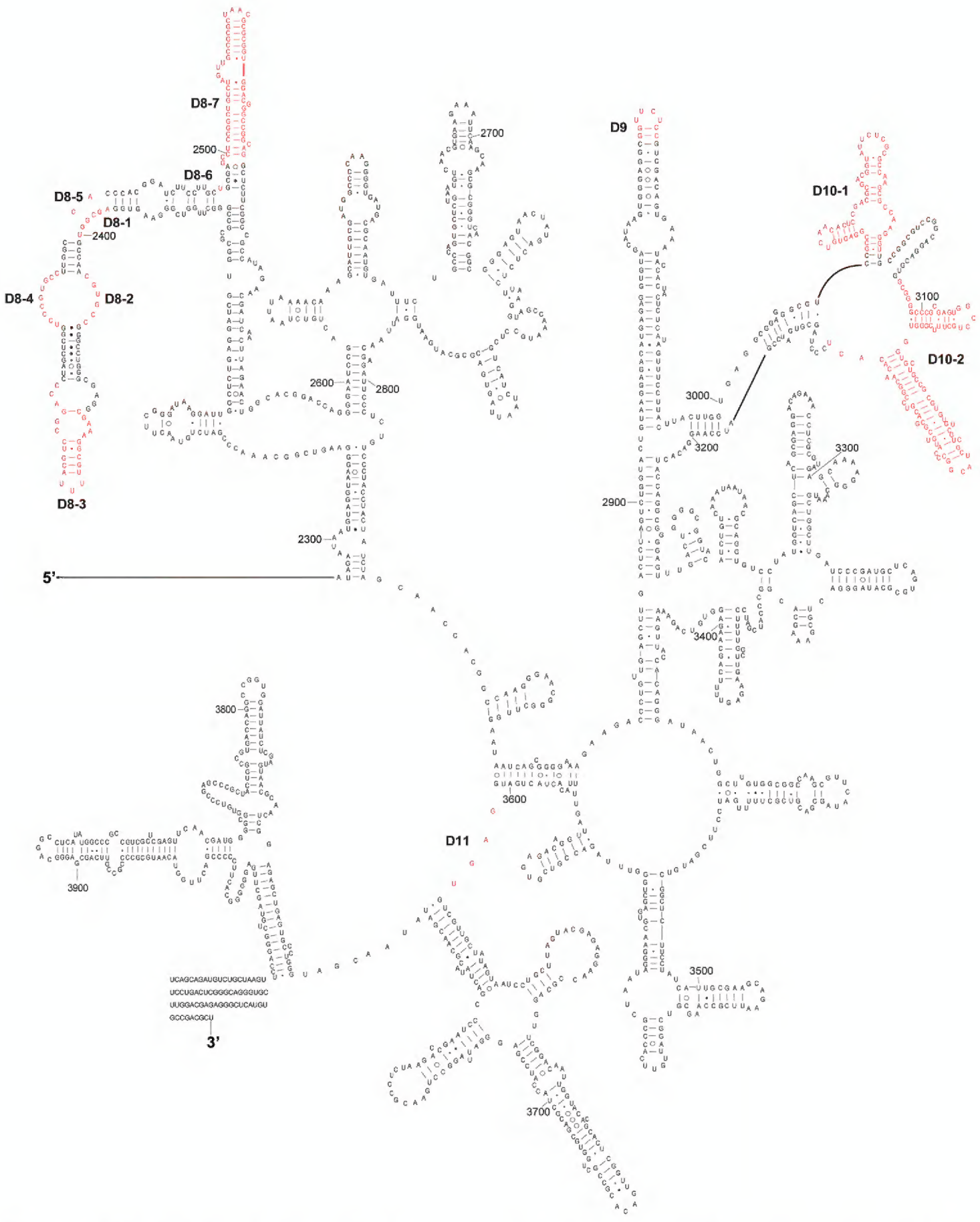

Figure 3. The 3'-half part of secondary structure model of 28S rRNA of *Eurydema maracandica*. The numbers D8 to D11 represent four LVRs.

The LVRs potentially serving as synapomorphies

Campbell et al. (1994) hypothesized that length expansion in 18S rDNA could serve as synapomorphy for Sternorrhyncha. Later, the concepts of "molecular morpho-metrics" (Billoud et al. 2000) and "morpho-molecular structures" (Ouvrard et al. 2000)

were introduced. It was found that length variation and indels could serve as synapomorphies for specific monophyletic groups of insects (Xie et al. 2009). The results of this study indicate that, the LVRs B, M, T, U and W of 18S rDNA, and D3-1, D3-2 and D3-3 of 28S rDNA could potentially serve as molecular synapomorphies (Table 1). According to the lengths of LVRs B and T of 18S rDNA, there only a few exceptions exist within Pentatomomorpha. Therefore, the 11nt length of LVR B and the 4nt length of LVR T very likely originate from the last ancestor of Pentatomomorpha and are the synapomorphies of this infraorder. In the same sense, the 10nt length of LVR M is a dominant state in the clade (Cimicomorpha+Pentatomomorpha), and the 9nt length of LVR U is very conservative within Neoheteroptera. Fig. 4 shows the secondary structure models of LVR W of Naboidea and Cimicoidea. Within Heteroptera, Naboidea and Cimicoidea have the same unique length of 5nt in this region. Taking the problem of the phylogeny within the superfamily Pentatomoidea for more instances, the 5nt length of LVR D3-1 and 19nt length of LVR D3-2 in 28S rRNA could possibly serve as synapomorphies for an Acanthosomatidae+Lestoniidae clade (Grazia et al. 2008). Additionally, the Paraphrynoveliidae+Macroveliidae clade shares the same length of LVR D3-1, D3-2 and D3-3. Gelastocoridae and Ochteridae have the same length of 4nt in D3-1 region. Therefore amplification of complete 18S and 28S rDNAs of more representatives very likely will result in more representative dataset for reconstruction of ancestral states and discoveries of more molecular synapomorphies for specific clades within Heteroptera.

Table 1. Monophyletic groups within Heteroptera with some LVRs serving as potential synapomorphies.

Monophyletic group	Reference	LVR (number of sequences examined)	
		18S rDNA	28S rDNA
Neoheteroptera	Štys et al. (1975) Schuh (1979) Wheeler et al. (1993)	9nt U (131)	
Paraphrynoveliidae + Macroveliidae	Damgaard (2008)		4nt D3-1 (3) 6nt D3-2 (3) 15nt D3-3 (3)
Gelastocoridae + Ochteridae	Hebsgaard (2004)		4nt D3-1 (4)
Cimicomorpha + Pentatomomorpha	Štys et al. (1975) Schuh (1979) Schaefer (1993) Wheeler et al. (1993)	10nt M (117)	
Naboidea + Cimicoidea	Schuh et al. (2009)	5nt W (36)	
Pentatomomorpha	Leston et al. (1954) Schuh (1979) Schaefer (1993) Wheeler et al. (1993)	11nt B (39) 4nt T (37)	
Acanthosomatidae+Lestoniidae	Grazia et al. (2008)		5nt D3-1 (8) 19nt D3-2 (8)

Note: Detailed information see Supplementary Files 4 and 5.

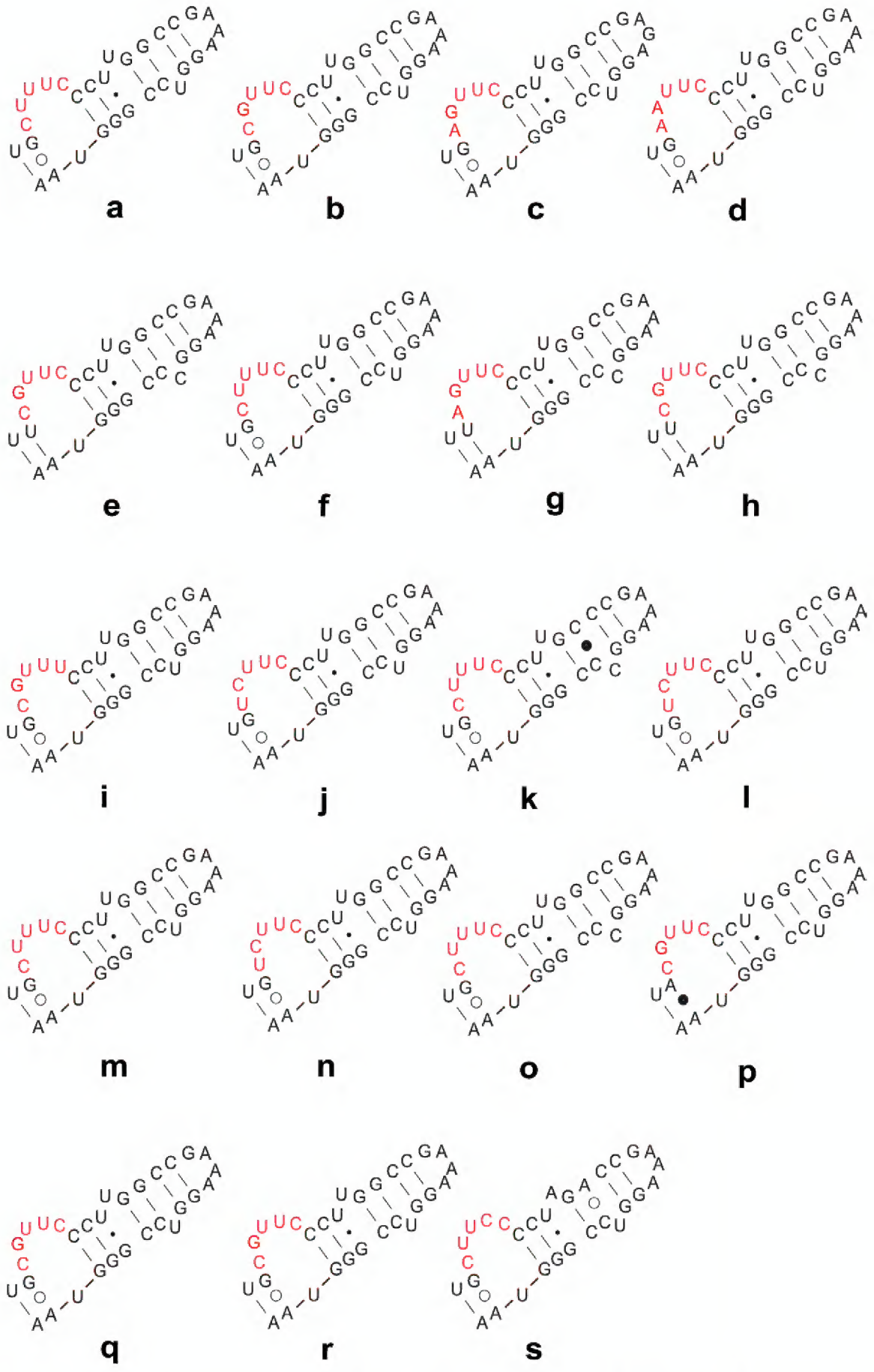

Figure 4. Secondary structure models of LVR W of Naboidea and Cimicoidea. These sequences are from 15 genera, 19 species of Naboidea and Cimicoidea. The species names and GenBank Accession numbers are as follow: Nabidae (**a**) *Nabis ferus* EF487300 (**b**) *Nabis flavomarginatus* GQ258424 (**c**) *Himacerus apterus* GQ258425; Lyctocoridae (**d**) *Lyctocoris beneficus* EF487298; Anthocoridae (**e**) *Anthocoris* sp. AY252319 (**f**) *Anthocoris confusus* EF487297 (**g**) *Anthocoris montanus* EF487307 (**h**) *Tetraphleps aterrimus* EF487295 (**i**) *Amphiareus obscuriceps* EF487301 (**j**) *Orius agilis* EF487296 (**k**) *Physopleurella armata* EF487308 (**l**) *Montandoniola moraguesi* EF487310 (**m**) *Xylocoris cerealis* GQ258395 (**n**) *Bilia* sp. GQ258406 (**o**) *Buchananiella crassicornis* GQ258407 (**p**) *Lasiochilus japonicus* GQ258410 (**q**) *Lasiochilus luceonotatus* GQ258411; Cimicidae (**r**) *Cimex lectularius* GQ258396; Curaliidae (**s**) *Curalium cronini* EU683128.

Acknowledgement

We are grateful to Dr. Petr Kment (National Museum, Prague) for confirming the identity of *E. maracandicum*. We also thank Dr. Zhong-Hua Fan (Nankai University) for helpful advices. This work was supported by the National Science Foundation of China (grant numbers 31222051, J1210005, J0930005, 30970350), China Postdoctoral Science Foundation (grant number 20110490769) and Science Foundation of Tianjin (grant number 11JCYBJC08100).

Reference

Barker SC, Whiting M, Johnson KP, Murrell A (2003) Phylogeny of the lice (Insecta, Phthiraptera) inferred from small subunit rRNA. Zoologica Scripta 32: 407–414. doi: 10.1046/j.1463-6409.2003.00120.x

Billoud B, Guerrucci MA, Masselot M, Deutsch JS (2000) Cirripede phylogeny using a novel approach: molecular morphometrics. Molecular Biology and Evolution 17: 1435–1445. doi: 10.1093/oxfordjournals.molbev.a026244

Calleros L, Panzera F, Bargues M, Monteiro F, Klisiowicz D, Zuriaga M, Mas-Coma S, Pérez R (2010) Systematics of *Mepraia* (Hemiptera-Reduviidae): Cytogenetic and molecular variation. Infection, Genetics and Evolution 10: 221–228. doi: 10.1016/j.meegid.2009.12.002

Campbell BC, Steffen-Campbell JD, Gill RJ (1994) Evolutionary origin of whiteflies (Hemiptera: Sternorrhyncha: Aleyrodidae) inferred from 18S rDNA sequences. Insect Molecular Biology 3: 73–88. doi: 10.1111/j.1365-2583.1994.tb00154.x

Damgaard J (2008) Phylogeny of the semiaquatic bugs (Hemiptera-Heteroptera, Gerromorpha). Insect Systematics and Evolution 39: 431–460. doi: 10.1163/187631208788784264

Damgaard J, Andersen NM, Meier R (2005) Combining molecular and morphological analyses of water strider phylogeny (Hemiptera-Heteroptera, Gerromorpha): effects of alignment and taxon sampling. Systematic Entomology 30: 289–309. doi: 10.1111/j.1365-3113.2004.00275.x

Ellis R, Sulston J, Coulson A (1986) The rDNA of *C. elegans*: sequence and structure. Nucleic Acids Research 14: 2345–2364. doi: 10.1093/nar/14.5.2345

Grazia J, Schuh RT, Wheeler WC (2008) Phylogenetic relationships of family groups in Pentatomoidea based on morphology and DNA sequences (Insecta: Heteroptera). Cladistics 24: 932–976. doi: 10.1111/j.1096-0031.2008.00224.x

Gutell RR, Lee JC, Cannone JJ (2002) The accuracy of ribosomal RNA comparative structure models. Current opinion in structural biology 12: 301–310. doi: 10.1016/S0959-440X(02)00339-1

Hall T (2011) Bioedit: a biological sequence alignment editor. v7.1.3. Ibis Therapeutics, a division of Isis Pharmaceuticals, Carlsbad.

Hassouna N, Michot B, Bachellerie JP (1984) The complete nucleotide sequence of mouse 28S rRNA gene. Implications for the process of size increase of the large subunit rRNA in higher eukaryotes. Nucleic Acids Research 12: 3563–3583. doi: 10.1093/nar/12.8.3563

Hebsgaard MB, Andersen NM, Damgaard J (2004) Phylogeny of the true water bugs (Nepomorpha: Hemiptera-Heteroptera) based on 16S and 28S rDNA and morphology. Systematic Entomology 29: 488–508. doi: 10.1111/j.0307-6970.2004.00254.x

Henry TJ (1997) Phylogenetic analysis of family groups within the infraorder Pentatomomorpha (Hemiptera: Heteroptera), with emphasis on the Lygaeoidea. Annals of the Entomological Society of America 90: 275–301.

Hillis DM, Dixon MT (1991) Ribosomal DNA: molecular evolution and phylogenetic inference. Quarterly Review of Biology 66: 411–453. doi: 10.1086/417338

Jung S, Lee S (2011) Molecular phylogeny of the plant bugs (Heteroptera: Miridae) and the evolution of feeding habits. Cladistics 28: 50–79. doi: 10.1111/j.1096-0031.2011.00365.x

Kjer KM, Carle FL, Litman J, Ware J (2006) A molecular phylogeny of Hexapoda. Arthropod Systematics and Phylogeny 64: 35–44.

Lalitha S (2000) Primer premier 5. Biotech Software and Internet Report: The Computer Software Journal for Scientists 1: 270–272. doi: 10.1089/152791600459894

Leston D, Pendergrast J, Southwood T (1954) Classification of the terrestrial Heteroptera (Geocorisae). Nature 174: 91. doi: 10.1038/174091b0

Marcilla A, Bargues MD, Ramsey JM, Magallon-Gastelum E, Salazar-Schettino PM, Abad-Franch F, Dujardin JP, Schofield CJ, Mas-Coma S (2001) The ITS-2 of the nuclear rDNA as a molecular marker for populations, species, and phylogenetic relationships in Triatominae (Hemiptera: Reduviidae), vectors of Chagas disease. Molecular Phylogenetics and Evolution 18: 136–142. doi: 10.1006/mpev.2000.0864

Neefs JM, Van de Peer Y, De Rijk P, Chapelle S, De Wachter R (1993) Compilation of small ribosomal subunit RNA structures. Nucleic Acids Research 21: 3025–3049. doi: 10.1093/nar/21.13.3025

Nonnaizab (1988) Fauna of Inner Mongolia (Hemiptera: Heteroptera). Volum [sic] I. Book 1. Inner Mongolia People's Publishing House, Hohhot, 2+2+469 pp. [in Chinese, English summary]

Oshanin B (1871) [On heteropterous insects of the Zaravshan Valley.] Izvestiya Imperatorskago Obshchestva Lyubitelei Estestvoznaniya, Antropologii i Etnografii pri Imperatorskom Moskovskom Universitete 8: 194–213 [in Russian]

Ouvrard D, Campbell BC, Bourgoin T, Chan KL (2000) 18S rRNA Secondary structure and phylogenetic position of Peloridiidae (Insecta, Hemiptera). Molecular Phylogenetics and Evolution 16:403–417. doi: 10.1006/mpev.2000.0797

Putshkov VG (1965) [Shield-bugs of Middle Asia (Hemiptera, Pentatomoidea).] Ilim, Frunze, 331 pp [in Russian]

Reineke A, Karlovsky P, Zebitz C (2002) Preparation and purification of DNA from insects for AFLP analysis. Insect Molecular Biology 7: 95–99. doi: 10.1046/j.1365-2583.1998.71048.x

Reuter JS, Mathews DH (2010) RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinformatics 11: 129. doi: 10.1186/1471-2105-11-129

Schaefer CW (1993) The Pentatomomorpha (Hemiptera: Heteroptera): an annotated outline of its systematic history. European Journal of Entomology 90: 105–105.

Schuh RT (1979) (Review of Cobben 1978). Systematic Zoology 28: 653–656. doi: 10.2307/2412577

Schuh RT, Slater JA (1995) True bugs of the world (Hemiptera: Heteroptera): classification and natural history. Cornell University Press, Ithaca, London, 336 pp.

Schuh RT, Weirauch C, Henry TJ, Halbert SE (2008) Curaliidae, a new family of Heteroptera (Insecta: Hemiptera) from the eastern United States. Annals of the Entomological Society of America 101: 20–29. doi: 10.1603/0013-8746(2008)101[20:CANFOH]2.0.CO;2

Schuh RT, Weirauch C, Wheeler WC (2009) Phylogenetic relationships within the Cimicomorpha (Hemiptera: Heteroptera): a total-evidence analysis. Systematic Entomology 34: 15–48. doi: 10.1111/j.1365-3113.2008.00436.x

Stichel W (1961) Pentatomomorpha, Pentatomidae, Trib. Eurydemini Dist. [partim]. In: Stichel W (Ed) Illustrierte Bestimmungstabellen der Wanzen. II. Europa. (Hemiptera-Heteroptera Europae) 4, Stichel, Berlin-Hermsdorf, (39–41): 609–656.

Štys P, Kerzíner I (1975) The rank and nomenclature of higher taxa in recent Heteroptera. Acta entomologica Bohemoslovaca 72: 65–79.

Tian Y, Zhu WB, Li M, Xie Q, Bu WJ (2008) Influence of data conflict and molecular phylogeny of major clades in Cimicomorphan true bugs (Insecta: Hemiptera: Heteroptera). Molecular Phylogenetics and Evolution 47: 581–597. doi: 10.1016/j.ympev.2008.01.034

Wang YH, Engel MS, Rafael JA, Dang K, Wu HY, Wang Y, Xie Q, Bu WJ (2013) A Unique Box in 28S rRNA is Shared by the Enigmatic Insect Order Zoraptera and Dictyoptera. PloS one 8: e53679. doi: 10.1371/journal.pone.0053679

Wheeler WC, Schuh RT, Bang R (1993) Cladistic relationships among higher groups of Heteroptera: congruence between morphological and molecular data sets. Insect Systematics and Evolution 24: 121–137. doi: 10.1163/187631293X00235

Xie Q, Bu WJ, Zheng LY (2005) The Bayesian phylogenetic analysis of the 18S rRNA sequences from the main lineages of Trichophora (Insecta: Heteroptera: Pentatomomorpha). Molecular Phylogenetics and Evolution 34: 448–451. doi: 10.1016/j.ympev.2004.10.015

Xie Q, Lin JZ, Qin Y, Zhou JF, Bu W (2011) Structural diversity of eukaryotic 18S rRNA and its impact on alignment and phylogenetic reconstruction. Protein and Cell 2: 161–170. doi: 10.1007/s13238-011-1017-2

Xie Q, Tian XX, Qin Y, Bu WJ (2009) Phylogenetic comparison of local length plasticity of the small subunit of nuclear rDNAs among all Hexapoda orders and the impact of hyper-length-variation on alignment. Molecular Phylogenetics and Evolution 50: 310–316. doi: 10.1016/j.ympev.2008.10.025

Xie Q, Tian Y, Zheng LY, Bu WJ (2008) 18S rRNA hyper-elongation and the phylogeny of Euhemiptera (Insecta: Hemiptera). Molecular Phylogenetics and Evolution 47: 463–471. doi: 10.1016/j.ympev.2008.01.024

Zheng LY (1983) Notes on genus *Eurydema* from China. Acta Scientiarum Naturalium Universitatis Nankaiensis 1983: 125–130 [in Chinese, English summary]

Appendix 1

Primer pairs and the annealing temperatures for amplifying rDNAs of *Eurydema maracandica*. (doi: 10.3897/zookeys.319.4178.app1) File format: Mircrosoft Word Document (doc).

Explanation note: List of primer pairs and annealing temperatures.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Citation: Yu S, Wang Y, Rédei D, Xie Q, Bu W (2013) Secondary structure models of 18S and 28S rRNAs of the true bugs based on complete rDNA sequences of *Eurydema maracandica* Oshanin, 1871 (Heteroptera, Pentatomidae). In: Popov A, Grozeva S, Simov N, Tasheva E (Eds) Advances in Hemipterology. ZooKeys 319: 363–377. doi: 10.3897/zookeys.319.4178.app1

Appendix 2

Alignment of 18S rDNA sequence of Heteroptera. (doi: 10.3897/zookeys.319.4178.app2) File format: DNA and Protein Sequence Alignment (fas).

Explanation note: The alignment of *Eurydema maracandica* 18S rDNA sequence with other 140 orthologous sequences of Heteroptera.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Citation: Yu S, Wang Y, Rédei D, Xie Q, Bu W (2013) Secondary structure models of 18S and 28S rRNAs of the true bugs based on complete rDNA sequences of *Eurydema maracandica* Oshanin, 1871 (Heteroptera, Pentatomidae). In: Popov A, Grozeva S, Simov N, Tasheva E (Eds) Advances in Hemipterology. ZooKeys 319: 363–377. doi: 10.3897/zookeys.319.4178.app2

Appendix 3

Alignment of D3 region in 28S rDNA sequence of Heteroptera. (doi: 10.3897/zookeys.319.4178.app3) File format: DNA and Protein Sequence Alignment (fas).

Explanation note: The alignment of *Eurydema maracandica* D3 region in 28S rDNA sequence with other 84 orthologous sequences of Heteroptera.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Citation: Yu S, Wang Y, Rédei D, Xie Q, Bu W (2013) Secondary structure models of 18S and 28S rRNAs of the true bugs based on complete rDNA sequences of *Eurydema maracandica* Oshanin, 1871 (Heteroptera, Pentatomidae). In: Popov A, Grozeva S, Simov N, Tasheva E (Eds) Advances in Hemipterology. ZooKeys 319: 363–377. doi: 10.3897/zookeys.319.4178.app3

Appendix 4

The summarization of length variation in 18S rDNA sequence of Heteroptera. (doi: 10.3897/zookeys.319.4178.app4) File format: Microsoft Excel Worksheed (xls).

Explanation note: The length of 12 LVRs in all currently known 18S rDNA sequences of Heteroptera.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Citation: Yu S, Wang Y, Rédei D, Xie Q, Bu W (2013) Secondary structure models of 18S and 28S rRNAs of the true bugs based on complete rDNA sequences of *Eurydema maracandica* Oshanin, 1871 (Heteroptera, Pentatomidae). In: Popov A, Grozeva S, Simov N, Tasheva E (Eds) Advances in Hemipterology. ZooKeys 319: 363–377. doi: 10.3897/zookeys.319.4178.app4

Appendix 5

The summarization of length variation in D3 region of 28S rDNA sequence of Heteroptera. (doi: 10.3897/zookeys.319.4178.app5) File format: Microsoft Excel Worksheed (xls).

Explanation note: The length of 3 LVRs in all currently known sequences of D3 region of 28S rDNA of Heteroptera.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Citation: Yu S, Wang Y, Rédei D, Xie Q, Bu W (2013) Secondary structure models of 18S and 28S rRNAs of the true bugs based on complete rDNA sequences of *Eurydema maracandica* Oshanin, 1871 (Heteroptera, Pentatomidae). In: Popov A, Grozeva S, Simov N, Tasheva E (Eds) Advances in Hemipterology. ZooKeys 319: 363–377. doi: 10.3897/zookeys.319.4178.app5